The León Antiphoner: MEI and Manual Encoding

Geert Maessen, Gregoriana Amsterdam, the Netherlands

In this paper we will illustrate the potential usefulness of traditional manually encoded data (MED) for supervision and verification of automated document transcription as prompted with the music encoding initiative (MEI). We will do so with reference to our MED for the early tenth-century León antiphoner, the most important witness for the lost chant of the mozarabic rite (Fernández de la Cuesta 2011), to be uploaded on GitHub in 2019. First we will sketch the historical context of MED and MEI. Then we will focus on different steps of automated document transcription with examples from the León manuscript.

Early examples of chant studies with MED include the edition of the eleven century Dijon tonary (Hansen 1979). Although the edition is still widely available, the data are lost. One of the earliest preserved sets of medieval chant was presented eighteen years later (Haas 1997). Several other sets have been presented since (e.g. Helsen 2008). Most of these sets used different encoding formats (often in ASCII text). Since the turn of the century efforts have been made to standardize formats for chant notation, including staffless neumes (Barton, Caldwell, & Jeavons 2005).

Since 2011 these efforts have merged into the much broader MEI, an international collaborative effort, based upon open standards, to capture the semantics of music notation documents in machine-readable data (Hankinson, Roland & Fujinaga 2011; Morent 2012). The strenght of MEI is best appreciated in the emerging contexts of optical music recognition, automated manuscript transcription and machine learning. MEI preludes the abandonment of manual transcription. As is well known, since 2013 the music encoding conference is the centre of MEI activities.

In August 2018 the Cantus Planus conference, the major forum in chant scholarship, was held in Växjö, Sweden. Here a panel of domain experts under direction of Ichiro Fujinaga presented the recent MEI description of basic neume forms of at least eight different manuscript traditions for medieval chant (with and without pitch information): Saint Gall, Laon, León, Square notation, South Germany, Benevento, Utrecht and Nonantola. To our knowledge, however, only five manuscripts have been completely encoded so far, including the two volumes of the Hartker antiphoner, written in adiastematic Saint Gall neumes (Helsen, Behrendt & Sexton 2012). For comparative computational analysis of these traditions, therefore, a lot of work remains to be done. Since this has, until recently, not resulted in many data sets, some scholars again produced traditional MED sets for their specific purposes. Recently e.g. a MED set of five medieval traditions with analysis has been presented (Van Kranenburg & Maessen 2017).

For automated manuscript transcription of the León antiphoner, at least eight steps should be formalized. Although most steps are necessary in all traditions, in each tradtition domain experts will remain necessary to guide the process. These experts in turn could greatly profit from independent traditional MED sets. With examples from the León antiphoner this paper will illustrate problems in these eight steps and point to the benefits of traditional MED for their solution:

1. *Define the basic neume shapes.* The presentation in Växjö suggested that MEI's involvement was still focused on this step. The León antiphoner, however, is much richer in its neume shapes than most other manuscript traditions. Simple neumes of less than four notes have many variants. The torculus (open-high-low) e.g. has over forty different forms. Classification of simple neumes is not always obvious. Since complex neumes may be considered combinations of simple neumes, here also classification problems do exist. We will give some examples of what may be listed as problematic cases, resulting from a

hypothetical computational comparison of MED with an automatic transcribed León manuscript.

2. *Recognize all different handwritten forms of each shape everywhere.* At least four different hands were distinguished by MEI's domain expert (De Luca 2016). The MEI encoding should not only be able to capture specific individual neume shapes from each scribe, but also to generalize above the variants within a single hand over the complete manuscript. Although MED sets supposedly will be less detailed to differences in scribal hands, they could obviously assist the domain expert to locate problems here.

3. *Recognize the right order of these neumes in all instances.* León neumes are not always written straight above the related text syllables. Sometimes they are mixed with neumes for next syllables. Sometimes strings of neumes are written bottom up in the margin. Sometimes neumes are apparently erased and replaced by new ones. Automated transcription will be in need of expert knowledge, and sometimes even research, to solve these problems. Here also, MED could be very useful to locate problematic passages.

4. *Recognize the mozarabic text script.* The script used for the lyrics in León is very different from the carolingian minuscal used in Saint Gall. Here, possibly MEI has to make an alliance with the original TEI (text encoding initiative) or encourage it to include mozarabic script. MED sets with included texts could help to ease this task.

5. *Split these texts into the right syllables.* Probably TEI and an appropriate dictionary could provide a solution here. MED might help.

6. *Associate the right neumes with the right syllables.* Although this is a general problem for neumatic notation, here also, domain experts will remain necessary. Again, for automated transcriptions free of errors MED could be helpful.

7. *Recognize the different rubrics and chant genres throughout the manuscript.* Even when most sophisticated machine learning techniques would be used, here also, TEI and domain experts will remain necessary. MED could help.

8. *Decide what are supposed to be individual chants.* This may seem a simple problem. However, many chants cross pages, or are completed at unexpected places. So, here again, the final word is to the expert, in assistance with MED.

In summary, traditional MED could be of great benefit for future MEI purposes; especially to trace problems in automated transcriptions. We planned to upload a traditional MED set of León chants on GitHub in summer 2019.

References;

Barton, L., Caldwell, J. & Jeavons, P. (2005). E-library of Medieval Chant Manuscript Transcriptions. In *Proceedings of the Annual Joint Conference on Digital Libraries (JCDL)*, (pp 320–39).

De Luca, E. (2016). Scribal Habits in the León Antiphoner. In *Cantus Planus, Study Group of the International Musicological Society, Papers read at the 18th meeting, Dublin, Ireland, August 2-7, 2016* (forthcoming).

Fernández de la Cuesta, I. (2011). *Liber Antiphonarium de toto anni circulo a festivitate sancti Aciscli usque ad finem. Librum Ikilani Abbati. Catedralis Ecclesiae Legionensis in Hispania codex signatus Nr. VIII.* online at: http://bvpb.mcu.es/es/catalogo_imagenes/grupo.cmd?path=26408

Haas, M. (1997). *Mündliche Überlieferung und altrömischer Choral, Historische und analytische computergestützte Untersuchungen.*

Hankinson, A., Roland, P. & Fujinaga, I. (2011). The Music Encoding Initiative as a Document-Encoding Framework. In *Proceedings of the 12th International Society for*

Music Information Retrieval Conference, ISMIR 2011, Miami, Florida, USA, October 24-28, 2011 (pp 293-298).

Hansen, F. E. (1979). *The Grammar of Gregorian Tonality, An Investigation Based on the Repertory in Codex H 159, Montpellier.*

Helsen, K. E. (2008). *The Great Responsories of the Divine Office, Aspects of Structure and Transmission.*

Helsen, K., Behrendt, I. & Sexton, A. (2012). "Venite et videte": First Results in the Optical Neume Recognition Project. In *Cantus Planus, Study Group of the International Musicological Society, Papers read at the 16th meeting, Vienna, Austria, August 21-27, 2011,* (pp 161-165).

Morent, S. J. (2012). Digital Editing of chant - tools in progress. In *Cantus Planus, Study Group of the International Musicological Society, Papers read at the 16th meeting, Vienna, Austria, August 21-27, 2011,* (pp 281-283).

Van Kranenburg, P. & Maessen, G. (2017). Comparing Offertory Melodies of Five Medieval Christian Chant Traditions. In *Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR 2017, Suzhou, China, October 23-27, 2017,* (pp 204-210).

Presentation needs:

Computer, projector and screen to display manuscript images (PDF / JPEG / PowerPoint).

Statement:

Our current interests related to music encoding primarily focus on computational analysis of medieval chant.